IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

DATA INTEGRITY AUDIT SCHEME BASED ON BLOCKCHAIN EXPANSION TECHNOLOGY

[1] Sadiya Mohammadi, [2] S. Jhansi, [3] Shaik Shaheen Kausar, [4] P. Swetha, [5] Dr. D. Aruna Kumari
[5] HOD, [4] Assistant Professor, [1,2,3] Student
Department of Computer Science Engineering, Vidya Jyothi Institute of Engineering and Technology, Hyderabad, India

ABSTRACT:

Increasing numbers of users are outsourcing data to the cloud, but data integrity is an important issue. Nowadays, security system is critical for ensuring privacy in many IT industries. In these industries, data tampering becomes one of the biggest security threats as the stored data can be modified in many unauthorized ways. This system is built to prevent data tampering to the user data using blockchain technology, called Data Integrity Audit Scheme based on Blockchain Expansion Technology. Due to the decentralization and immutability of blockchain, more and more researchers tend to use blockchain to replace third-party auditors. To provide integrity to the data, this system used an algorithm called SHA algorithm which used the generated hash codes to validate the data. Data in this system are generally the criminal reports of the people. To ensure data security, a reward pool mechanism is introduced. Comprehensive analysis from aspects such as storage, batch auditing and data consistency proves the correctness of the scheme. Experiments on the Ethereum blockchain platform demonstrate that this scheme can effectively reduce storage and computational overhead. Data integrity usually means maintaining the data at the original stat over its entire life-cycle.

Keywords: NodeJS, Express Server, JavaScript and HTML/CSS/Bootstrap programming languages.

I. Introduction

There are many technologies or methods available nowadays that can strengthen the security of the user data efficiently and effectively. One of the technologies is called Blockchain. It is basically a distributed database or a public ledger that stores valuable data in a secured and tamper-proof way. It has many blocks connected each other and forms like a chain. Each block contains a cryptographic hash of the previous block, a timestamp and the user data. In blockchain, data are stored in many different distributed network nodes or devices, which are referred to users. The users are connected to form a blockchain network and have the responsibility to maintain the blockchain. The security of the blockchain is relied on an algorithm called Proof-Of-Work algorithm where a nonce is increasing its values until a block hash code with leading zero bit is calculated. Despite the great success of cloud storage, it also faces various challenges, and its security, reliability and privacy have always been a serious issue. After the user stores the data on the cloud server, the server provider may damage or delete the user data due to various factors, verifying the integrity of outsourced data becomes a crucial issue in cloud storage. Remote data integrity audit technology is very convenient and safe to help users check the integrity of data stored in outsourced.
Users can be connected each other to form a blockchain network. Users in the same network can add the data to the blockchain and the data can also be reviewed by all users. Blockchain are distributed into all connected users. Data tampering is proved by the way that if any connected user changes the data in the blockchain, the data added by that user will not be accepted by other honest users. Data integrity is ensured by the way that the honest users who don't change the blockchain's data will be able to add the data to blockchain. Honest users mean the users who didn't change any data in the blockchain.

The paper presents a dynamic system that uses blockchain technology to store criminal data, it uses the following technology to build:

1. SHA-256: It stands for Secure Hash Algorithm which is the cryptographic hash functions for computing a condensed digital representation or a hash code. It includes five algorithms which are SHA-1, SHA-224, SHA-256, SHA-384 and SHA-512. The later four algorithms are referred as SHA-2 version.
2. Node.js: The server-side of the system is built using Node.js. It is a JavaScript runtime that allows developers to build scalable and efficient web applications.
3. Express.js: Express is a minimal and flexible Node.js web application framework. It's used for the development of the back end and API of our application.
4. HTML & CSS: The client-side of the system is built using HTML, CSS and JavaScript. It is responsible for rendering the web pages and handling the user interaction and communicates with the server-side using HTTP requests.
5.

II. Literature review

There are several related works available which has been already published. In this section we will analyse the related surveys approaches for the problems and their solutions and extend it to make application. Below are some literature reviews:

"A blockchain-based flexible data auditing scheme for the cloud service" by F. Kefeng, L. Fei, Y. Haiyang, and Y. Zhen. They have designed a blockchain-based flexible cloud data auditing scheme. In this scheme, a decentralized auditing framework is proposed to eliminate the dependency on the third-party auditor, which increases the stability, security and performance of the whole scheme.

Since the cloud service provider can automatically generates auditing proofs, our scheme can relieve the communication burdens of the cloud service provider. The proposed scheme also adapts the Merkle Hash tree to improve the verification performance.
"A blockchain-based flexible data auditing scheme for the cloud service" by F. Kefeng, L. Fei, Y. Haiyang, and Y. Zhen. They have designed a blockchain-based flexible cloud data auditing scheme. In this scheme, a decentralized auditing framework is proposed to eliminate the dependency on the third-party auditor, which increases the stability, security and performance of the whole scheme. Since the cloud service provider can automatically generates auditing proofs, our scheme can relieve the communication burdens of the cloud service provider. The proposed scheme also adapts the Merkle Hash tree to improve the verification performance.
"A blockchain based data integrity verification for cloud storage with T-Merkle tree" by K. He, J. Shi, C. Huang, and X. Hu. They have designed a blockchain based integrity verification scheme for large-scale cloud data using T- Merkle Tree. In this, data tags are generated by ZSS short signature and stored on blockchain, and a new verification method based on ZSS short signature is proposed.
"Identity-based public data integrity verification scheme in cloud storage system via blockchain" by Y. Yuan, J. Zhang, W. Xu, and Z. Li. in this paper, we provide a novel idea to construct the integrity verification scheme via blockchain. The construction of the proposed scheme is based on identity-based encryption (IBE) which avoids the complex certificate management caused by the public key infrastructure (PKI).
"A blockchain-based cloud data integrity verification scheme with high efficiency" by G. Xie, Y. Liu, G. Xin, and Q. Yang. This paper improves some defects of the previous methods and proposes an efficient cloud data integrity verification scheme based on blockchain. This paper has proposed a lattice signature algorithm to resist quantum computing and introduced cuckoo filter to simplify the computational overhead of the user verification phase.
"Data tag replacement algorithm for data integrity verification in cloud storage" by G. Xu, S. Han, Y. Bai, X. Feng, and Y. Gan. A secure and reliable data tag replacement algorithm DTRA. The proposed algorithm uses proactive and reactive replacement to update the data owner's private key.
"Auditable attribute-based data access control using blockchain in cloud storage" by A. V. Ezhil, G. K. Indra, and K. Kulothungan. A new data sharing system auditable attribute-based encryption scheme that integrates the advantages of blockchain technology with attribute-based access control. A designed trustworthy scheme which uses blockchain to provide attribute-based secure data sharing with integrity auditing. It also provides compensation to data owners, if their data integrity is lost.

III. Proposed methodology

This paper represents a Data Integrity Audit Scheme Based on Block Chain Expansion Technology using modern technologies with a step-by-step methodology as discussed below.

a) Concept & Design: A blockchain is a growing list of shared and immutable distributed ledgers that facilitates the process of recording transactions and tracking recorded history among blockchain network. Virtually anything of value can be tracked back. Blocks in the blockchain are linked with cryptography that each block contains a cryptographic hash of the previous block, a timestamp and data. Security of blockchain is designed to be resistant to modification of the recorded data. Interfering with transactions on the blockchain is extremely difficult due to the complex cryptography employed and its nature of distributed ledger.

1) Block header: The block header consists of three sets of block metadata as shown in Figure 1. Metadata is data that provides information about other data. Firstly, there is a reference to a previous block hash, which connects this block to the previous block, lying in the blockchain. The second set of metadata relates to the mining competition; namely the difficulty, timestamp and nonce. Lastly, the third piece of metadata is the Merkle Tree root; a data structure used to summarize all the transactions in the block in an efficient manner.
2) Genesis Block: A genesis block is the first block of a blockchain. It has a block number of 0. The genesis block is almost hardcoded in the blockchain as it is a special block which does not contain its previous block.
3) Decentralized consensus: One of the most exciting aspects of blockchain technology is that it is entirely decentralized. The decentralized nature of blockchain technology means that it doesn't rely on a central point of control. Elimination to a single authority makes the system considerably more secure and safety. To securely transact the data without relying on central authority, blockchain utilizes an innovative protocol, called consensus protocol, across a network of nodes to validate and record blockchain data in an incorruptible way.

4) Nodes: A node is a device on a blockchain network that allows the blockchain to function and survive. Nodes are distributed across over the networks and have the responsibility to maintain the blockchain.

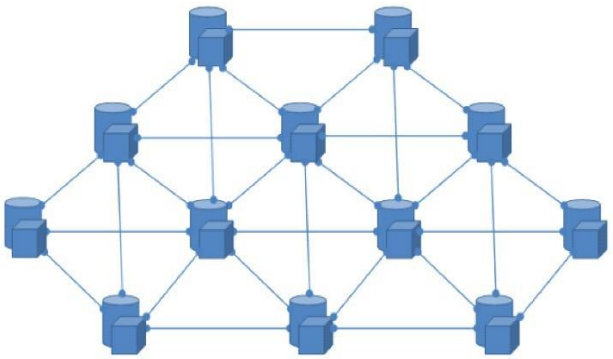

Figure 1: Decentralized networks

b) Implementation & Method: The system design is implemented as data integrity on blockchain. The system designs show that how a new block is created after the users added the data, how data tampering is proved by using data validation, and how to validate a new block that is broadcasted by other users in the network.

1) System Design for Creating New Block: After the data is ready, SHA-256 algorithm is applied to the hash code of the previous block and the data. The process is looking for a none value that when hashed with SHA-256, it results a satisfied hash code of the current block, starting with four zeros. After finding the satisfied hash code, a new block is created with the current timestamp. The data, nonce value, hash code of the previous block and hash code of the current block are then added to the new block.
2) System Design for Data Tampering: Blocks in the blockchain are linking each other using cryptographic hash functions. As each block contains a hash code of the previous block and that of the current block, data validation is achieved by comparing the hash codes of the two blocks.
3) System Design for Validating New Block: Validation for the blocks required because only the valid blocks will be added to the blockchain. Data validation to the block is done by comparing the current hash code of the block and the previous hash code of the new block.

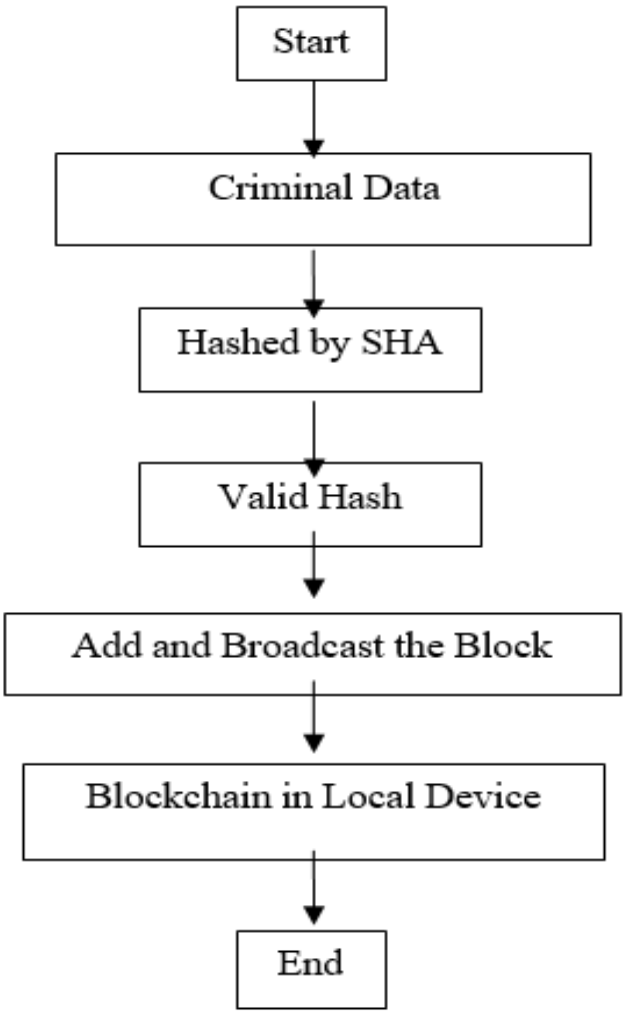

Figure 2: System design for creating new block.

IV. Results and discussions

A detailed guide is given below about how the Data Integrity Audit Scheme Based on Block chain Expansion works for different users and different purposes.

Step 1: When the users start running the system, this is the first page that will be shown. The system is greeting with the text "Data Integrity on Blockchain" to the user.

Step 2: This is the input page after the users have clicked the second button on the top bar. There are eight input fields that are the case number, the reason of charge, the name of the person, his/her gender, his/her social security number, date of birth, the date that the person offended a criminal and the date that the case is disposed. The input fields and the create button are initially shown red because no data is filled in the fields.

Figure 3: Home Page

Step 3: This is the page that shows all the blockchain data added by all the network devices. There is a text that shows “Data in the Blockchain” and four colorful buttons which are to refresh the blockchain data, to check the blockchain integrity, to synchronize the blockchain data if a new node is connected for the first time and to insert the modified data to the blockchain. Below the buttons, there are blocks that show the blockchain data visually. The block will increase itself dynamically and the number of blocks depends on the data added by the users to blockchain. Users can slide from left to right to see the successive block list.

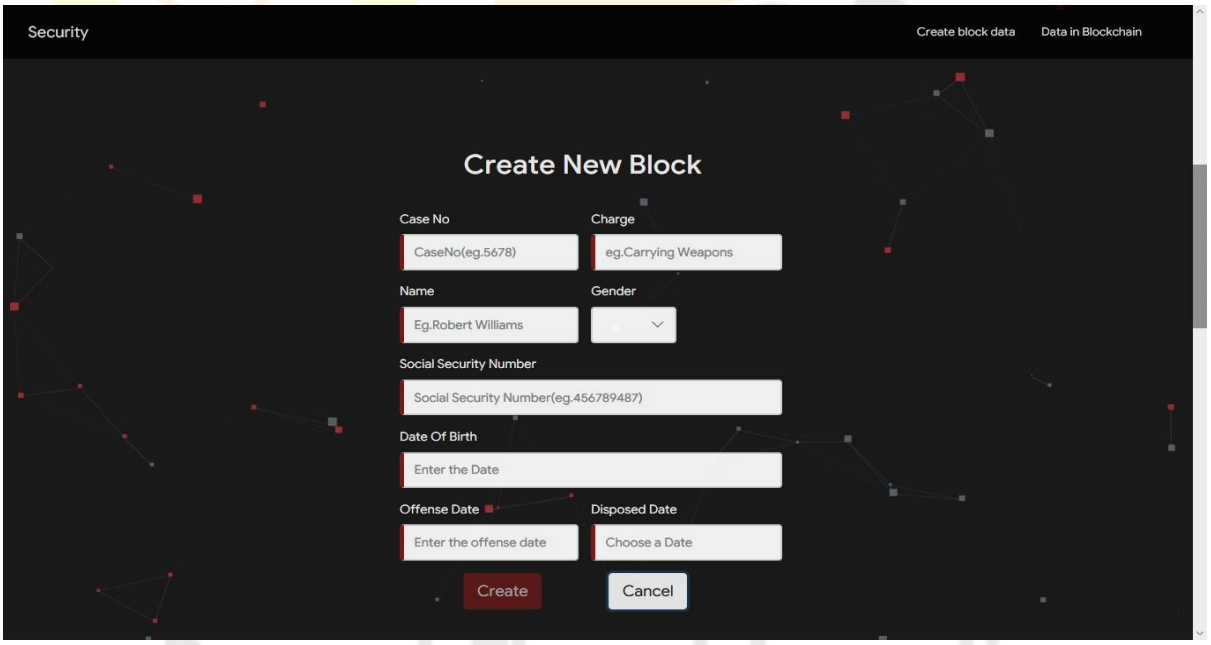

Figure 4: Input Page

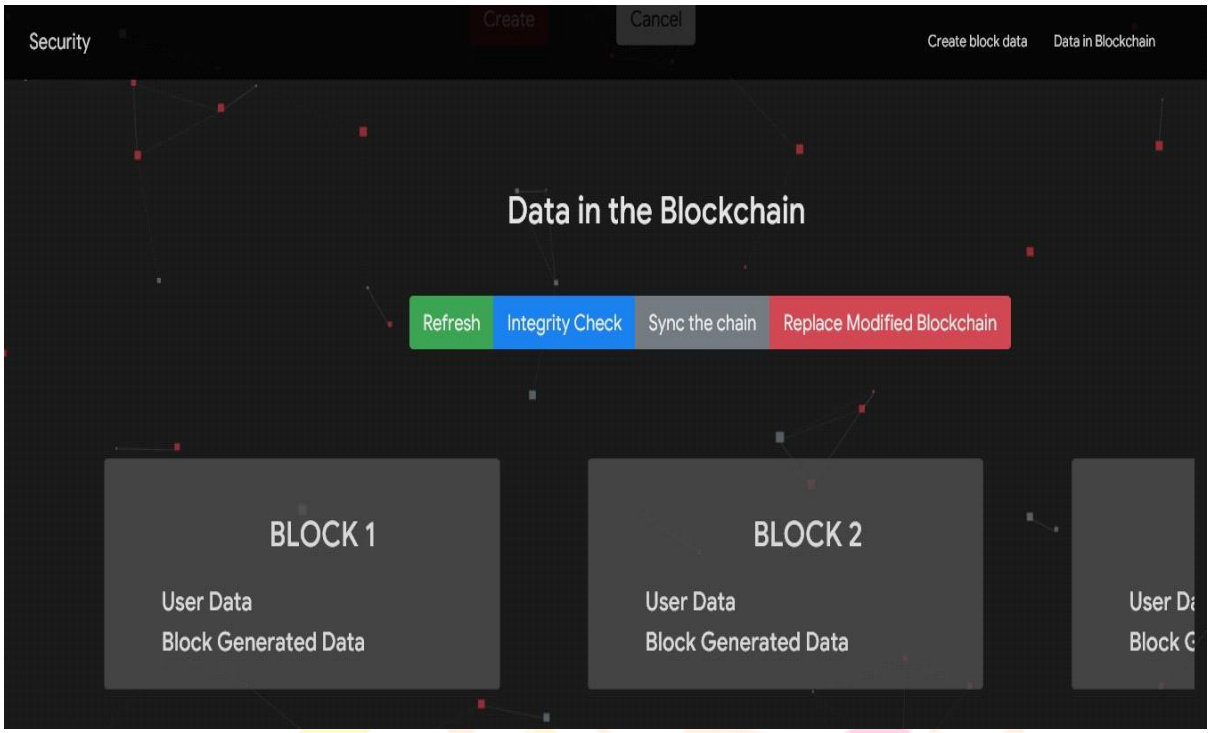

Figure 5: Blockchain Page

Step 4: The first block created by the users starts at the block number two and is adjacent to the genesis block. The new blocks created by the users will be added dynamically from the right side. In block 2, the user data are the values stored in the blockchain.

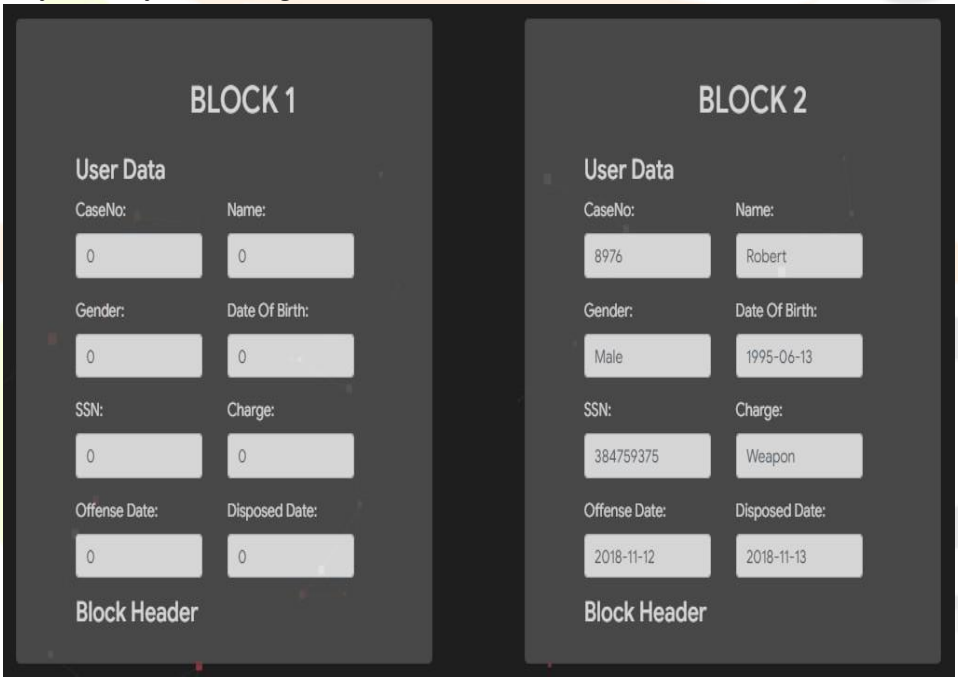

Figure 6: User Data

Step 5: Integrity check button is to notify the users about data in the blocks are correct. This is done by comparing the current hash code of the previous block and the previous hash code of the current block. The blocks will show in green color if the hash codes are the same. The comparing process take place from genesis block to block N. The blocks with green color means that the data in the blocks have integrity.

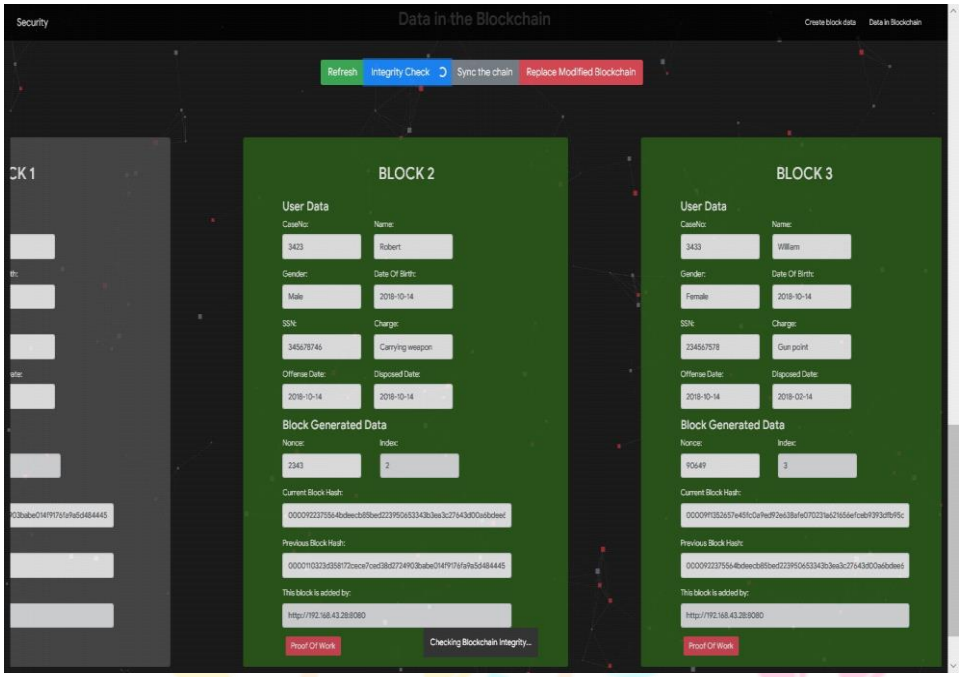

Figure 7: Checking Integrity on Block Data

Step 6: In this system, users can attempt to tamper the data in the blockchain. The data in the block 3 is modified by user and it lead to the modification to the rest of the block. “Replace Modified Blockchain” button will replace the current modified blockchain with the original one. After doing it, original blockchain cannot be recovered without restarting the node again. It will lead the node to be an invalid one in the network. The node cannot synchronize the latest blockchain with other users.

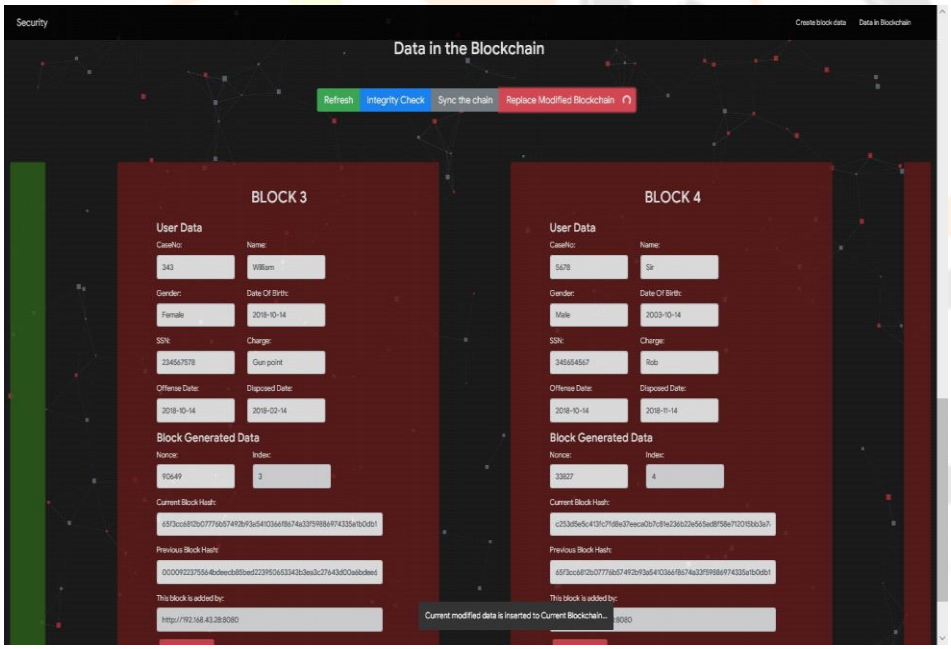

Figure 8: Modifying Current Blockchain

V. Conclusions

This system is implemented to demonstrate how a blockchain works. It is useful for the beginners who are curious about the blockchain technology. They can understand it clearly by using the system. It contains all the minimal functionalities of a blockchain and coded in easy way. The system provides an easy-to-use user interface that allows the users to perform data validation, data storing and data tampering. The user interface contains the visual effects that help users understand more quickly. In this system, users can connect each other and form a blockchain network. By using the system, user can have the knowledge of understanding proof-of-work algorithm and more and more. One of the trended technologies is the blockchain. This technology is widely used in many companies because it secures the information effectively and efficiently. This article proposes a data integrity scheme based on block chain expansion technology. In our scheme, we use the blockchain network to overcome some of the shortcomings of traditional auditing, improving the efficiency and security of the scheme. Data tampering is prevented by data validation using SHA algorithm. The nodes can be connected to form a blockchain network. The node with modified blockchain will not be considered as a valid node. The data added by that node will not be accepted by other nodes.

References

[1] **Investopedia, www.investopedia.com/terms/d/distributed-ledgers.asp, 2018.**
[2] **Mr. Vijay Divecha, Data Integrity in view of the Manufacturer,2018.**
[3] **FIPS PUB, Descriptions of SHA-256, SHA-384, and SHA-512, 2018.**
[4]**Ericsson,www.ericsson.com/en/security/data-centric-security/blockchain-data- integrity, 2018.**
[5] **Satoshi Nakamoto, Bitcoin (a peer-to-peer cash system), 2008.**